DOCUMENT RESUME

ED 373 099 TM 021 994

AUTHOR	Macintosh, Henry G.
TITLE	A Comparative Study of Current Theories and Practices in Assessing Students' Achievements at Primary and Secondary Level. Final Report. IBE Document Series, Number 4.
INSTITUTION	International Association for Educational Assessment.
SPONS AGENCY	International Bureau of Education, Geneva (Switzerland).
PUB DATE	Apr 94
NOTE	56p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Comparative Education; *Educational Assessment; Educational Diagnosis; Educational Practices; Educational Theories; Elementary Secondary Education; *Foreign Countries; Questionnaires; Selection; Test Results; *Test Use; *Theory Practice Relationship
IDENTIFIERS	Australia; Bahrain; England; Guatemala; International Surveys; Israel; Malaysia; Namibia; Poland; Scotland; Slovenia; Wales

ABSTRACT

A study was conducted to examine and compare, in 10 selected countries, the interaction between current conceptions and theories of assessment and assessment practices. Of particular interest was the nature of changing practice and theory in assessing the learning achievements of students in elementary and secondary education, as well as the sociocultural and technical factors that are associated with gaps between theory and practice. Countries participating were Australia, Bahrain, England and Wales, Guatemala, Israel, Malaysia, Namibia, Poland, Scotland, and Slovenia. A designated representative of an institution specializing in assessment in each country completed a questionnaire about assessment practice and theory. Across all participating nations, assessment appears to be more widely used for selection than for diagnosis. Most assessment is restricted in scope and operates in a relatively restructured curriculum. In addition, there is little pressure to change in these countries. Some reasons for the situation are explored. Two appendixes list the countries and organizations and contain the questionnaire. (Contains 41 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 373 099

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
F. NACEREDDINE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A COMPARATIVE STUDY OF CURRENT THEORIES AND PRACTICES IN ASSESSING STUDENTS' ACHIEVEMENTS AT PRIMARY AND SECONDARY LEVEL

FINAL REPORT

HENRY G MACINTOSH

International Association for Eductional Assessment

Prepared for the
International Bureau of Education
PO Box 199, 1211 GENEVA 20, Switzerland
January 1994

TM 021 994

IBE Documents Series. Number 4. April 1994

A COMPARATIVE STUDY OF CURRENT THEORIES AND PRACTICES IN ASSESSING STUDENTS' ACHIEVEMENTS AT PRIMARY AND SECONDARY LEVEL

FINAL REPORT

HENRY G MACINTOSH

International Association for Eductional Assessment

Prepared for the
International Bureau of Education
PO Box 199, 1211 GENEVA 20, Switzerland
January 1994

CONTENTS

	Page
Introduction	1
CHAPTER I: The Curriculum and its Assessment	6
CHAPTER II: Assessment: Policy, Practice and Purpose	21
Conclusion	32
References and Bibliography	35
Appendix A - List of Participants	38
Appendix B - Questionnaire	40

INTRODUCTION

Following discussions between the two bodies, the Executive Committee of the International Association for Educational Assessment (IAEA) agreed in January 1992 to undertake, on behalf of the International Bureau of Education (IBE) in Geneva A Comparative Study of Current Theories and Practices in Assessing Students' Achievements at the Primary and Secondary Level of Education. IAEA appointed Henry Macintosh, currently its Treasurer, to manage the Project on its behalf.

OBJECTIVES OF THE STUDY

The aim of the Study was to examine and compare in ten selected countries the interaction between current conceptions and theories of assessment[1] and assessment practices.

In particular the Study sought to identify in a comparative context:

(a) The nature of changing practices and theories during the past twenty to twenty-five years in assessing the learning achievements of students in primary and secondary education.

(b) The socio-cultural and technical factors associated with the gaps between assessment theory and assessment practice.

(c) Ways to reinforce links between theory and practice by increasing the relevance of the process of assessment in both local and national socio-cultural environments and the needs of individuals.

METHODOLOGY

The following ten countries were selected jointly by IBE and IAEA to participate in the Study: Australia; Bahrain; England and Wales; Guatemala; Israel; Malaysia; Namibia; Poland; Scotland and Slovenia and all agreed to take part. In each country an institution which specialized in assessment was identified (seven of these were members of IAEA and the others were identified by IBE). A designated individual within that institute was given responsibility for the in-country study, but was asked to consult as widely as possible and when replying to indicate the names and positions of those spoken or written to.

---

[1] The term assessment is used throughout in the Study and in this report to include all forms of testing or examination of students' intellectual aptitudes, abilities and knowledge

A full list of the participating institutions and the individuals involved is attached as Appendix A. It should be borne in mind that the numbers involved in the study were small; the largest number of contributors being seventeen in Guatemala. In consequence an individual's own interests could colour the responses made, particularly when the scope of the Study did not permit much in the way of subsequent clarification.

For the purpose of collecting the in-country information a three-section questionnaire was designed in both English and French. A copy of the English version of the questionnaire is attached as Appendix B.

The first section asked for information about the nature of the school curriculum and assessment under four broad headings: the curriculum; statutory assessment which is legally required in schools; non-statutory formal assessment and classroom assessment. The questions were largely closed in nature and were designed to elicit factual information.

The second section invited an evaluation of current developments in assessment policies and practices within the education system of the country concerned. Although respondents were asked to direct their replies to specific questions there was plenty of opportunity for expression of opinions.

The third section invited the respondents to address any general issues which they felt had not been covered in the questionnaire. Some guidance as to possible issues was provided but they were encouraged to address any topic which they believed to be important or relevant.

When replying to the questionnaire, whose terms were defined wherever it was thought differences of interpretation might arise, respondents were asked to:

(a) concentrate solely on assessment theory and practice in relation to students during the years of compulsory schooling and to exclude any developments which affected students in further, higher or vocational education;

(b) consider the full range of assessment practices, formal and informal, as they existed in their countries in primary and secondary education: for example - large-scale public examinations, the use of standardised tests, testing for selection or other purposes, procedures for marking and grading, and everyday questioning and assessment strategies adopted in classrooms as a regular part of teaching and learning;

(c) indicate where particular issues did not apply to their country;

(d) use the section and sub-section numbers set out in the questionnaire when structuring their responses in order to facilitate later analysis.

The responses to the first two sections were to be provided twice, once with regard to primary education and once with regard to the compulsory years of secondary education. For the purposes of the study compulsory schooling was assumed to consist of primary education starting at or about the age of five and continuing until the age of eleven and secondary education for students aged 11-16. In countries where these definitions did not apply respondents were asked to describe the structures that did exist.

A trial version of the questionnaire was completed informally by one of the participating countries (Australia) and amended in the light of comments received. The final version was issued to those taking part in May 1992 and they were asked to indicate problems or points which required clarification before starting work. No queries were received. The original completion date (December 1992) was extended to March 1993 by which time all ten questionnaires had been returned.

THE PREPARATION OF THE FINAL REPORT

As indicated earlier the aims of the partly-closed and a partly-open questionnaire were threefold:

i) to ensure that assessment was placed within the context of curriculum in schools;

ii) to encourage consideration of and opinions about assessment practices, formal and informal, statutory and non-statutory, radical and main stream (as defined by the respondents themselves) at the present time and in relation to changes that had taken place over the previous twenty to twenty-five years;

iii) to invite an overview on broader issues within specific countries; for example how far assessment practice is/was determined by socio-cultural factors; how far assessment practice is/was constrained by limitations in thinking on matters such as the range of purposes which assessment should serve.

There are obvious difficulties inherent in this approach. For example:

ensuring that a common terminology exists - the term assessment itself means many different things to different people;

making comparisons between countries with strongly centralised systems (England and Wales) and those like Australia where there is a multiplicity of state systems;

generalising between countries which are essentially monocultural (for example, Bahrain) and those which are multicultural (for example, Israel).

Such concerns of course challenged the respondents to the questionnaire as much as they challenged those responsible for pulling the threads together and providing an overview. To this end a broad outline framework had first to be devised. After considerable discussion it was decided to base this upon the answers to eight questions which were asked of each of the ten completed questionnaires.

i) The extent to which assessment practices, both past and present, have been determined by, influenced by, or linked to a national curriculum whether legally enforced or de facto.

ii) The nature, extent and purpose(s) of statutory assessment whether legally enforced or existing in quasi-legal forms.

iii) The role, extent and nature of informal assessment.

iv) The extent and degree of moves towards simplicity in assessment and the reasons for this.

v) The existence of a wide or narrow range of purposes for assessment, what they were and the relative importance attached to them.

vi) The status (high or low) enjoyed by different forms of assessment and what these were.

vii) The characteristics of the assessment procedures in use and the degree to which a wide or narrow range of procedures was used:

(a) in formal statutory assessment,
(b) in informal classroom practice.

viii) The nature and degree of change that had taken place over the past twenty to twenty-five years and the reasons for this.

It was not of course possible to supply answers to all these questions and there were inevitably differences in the way in which words like `simplicity' and `informal' were interpreted; although it was hoped that the questions themselves would clarify this and would also draw attention to possible other questions which might need to be asked and answered.

These questions and the answers provided the framework for the report which follows. It is a partly descriptive and partly evaluative look at the relative impact of social, cultural, economic and political factors upon current and past assessment practice within the participating countries.

ACKNOWLEDGEMENTS

The writer would like to thank Juan Carlos Tedesco, the Director of IBE; George Thomé, the former Director of IBE, Dr Barry McGaw and Dr Francis Ottobre, President and Executive Secretary respectively of IAEA for their support and Tosio Ohsako, Programme Specialist at IBE who drew up the original proposal and selected the participating countries. This report would not of course have been written without the hard work of those who completed the questionnaires. Their interest and support has been much appreciated.

CHAPTER I

THE CURRICULUM AND ITS ASSESSMENT

THE RESPONSES

(i) Australia

Although many aspects of organisation and educational delivery are common to the eight State education systems within Australia, each differs in some respects from all the others.

Primary school students follow a largely common curriculum, most usually: English or Language Arts, Mathematics, General Science, Social Science or Social Studies, Health and Physical Education, Art. Depending on such factors as the site, availability of staff and resources schools may also offer a range of optional subjects; for example: Music, Computer Studies, Languages other than English (LOTE) and Health and Human Relationships. Religious Studies are often compulsory in Catholic schools and church-run Independent schools.

Students in secondary schools also follow a largely common curriculum during the years of compulsory schooling, although some elements may be alternatives in various years. The range of subjects is essentially the same as in primary schools. Assessments are made and reported within all of a school's offerings, with the possible exception of sport.

Broadly speaking, the curriculum for State schools is determined by government regulation, and for Catholic schools by regulations issued by local Catholic Education Commissions. The extent to which there is central control over provision, for example in the form of an official syllabus which must be followed in detail, varies from State to State. Most offer some form of curriculum framework which allows schools a good deal of autonomy. Only in one state, which has in place a fairly prescriptive 'Unit Curriculum', is central control over curriculum provision relatively tight. Independent schools are free to determine their own curriculum within general guidelines issued by government boards. In practice the majority follow a curriculum broadly the same as that in state schools.

The needs of students with learning difficulties are generally met in mainstream classes. Increasingly this is also the case for those who have special disabilities, with separate provision only for students with severe physical or mental handicaps. Schools may develop extension or enrichment programmes to meet the needs of students for whom English is not their native or home language.

Where there have been curricular changes, generally they have not been formally imposed. Rather they have been developed by schools as an extension of their own practice or as a result of the efforts of subject associations. There have been significant thrusts in several areas, notably: Science, Oral Language, Health and Human Relationships, Computer Education. Media Education, Performance Arts and Environmental Science have also - albeit to a lesser extent - entered mainstream educational practice.

In most of the Australian states, schools are largely free to develop and implement their own assessment policies unencumbered by statutory requirements. There are four exceptions:

> New South Wales: there is a 'Basic Skills Testing Program' under which all students in Years 3 and 6 are tested in basic literacy (reading and editing skills) and numeracy, using common instruments.
>
> Western Australia: a program 'Monitoring Standards in Education' is being mounted for Years 3, 6 and 10. This will involve testing in a random sample of schools to check the performance of students - also in basic literacy and numeracy.
>
> Northern Territory: samples of students' work in Year 10 in a range of subjects are moderated annually as a check on standards prior to the award of the State's school certificate.
>
> Victoria: students' performances in literacy, numeracy, Science and Social Science are randomly sampled; one such area each year on a rolling cycle.

The most developed of these, the Basic Skills Testing Program, was introduced for the purpose of providing a longitudinal view of standards achievement in the target areas. It involves the testing of all students at the stipulated grades, within the exception of those who have immigrated within the previous twelve months. Draft test items are scrutinised before use for possible socio-economic bias. Trial test item statistics are also inspected to check for evidence of any such bias.

The only modes of assessment used are multiple choice and modified cloze exercises. Assessment results are criterion-referenced. The scales are constructed so that it is possible to make comparisons across time, between levels and between schools. Individual scores are compared with a profile of item performance characteristics across the whole cohort and reported to schools and parents in such a way that these comparisons are easily understood.

(ii) **Bahrain**

Formal education started in 1919 with the opening of the first primary school for boys. The first primary school for girls opened in 1928. Primary schooling (6 years) and intermediate schooling (3 years) jointly provide basic education. Secondary schooling lasts for a further three years.

State schools are not co-educational. All students follow a compulsory common curriculum throughout the six grades of primary education: Religious Education, Arabic, English (from the fourth grade), Mathematics, Science and Technology, Social Studies, Fine Arts, Physical Education, Music. In addition, girls' schools offer Family Education.

At the intermediate level also, students follow a common curriculum. It is broadly similar to the primary curriculum, with the addition of Practical Studies.

The Ministry of Education is responsible for ratifying and supervising the curriculum offered in state schools. In private schools the curriculum is determined partly by government legislation and partly by the school. Students with mental or physical disabilities may attend special schools

which are mostly financed and directed by private agencies. But, as yet, most schools offer only limited support to gifted students, to students for whom Arabic is not their native language, or to students with learning difficulties.

During the past 20-25 years there have been many changes, largely directed towards: greater integration of subjects, some reduction in the extent to which boys and girls followed different curricula, new approaches to pedagogy and the revision of textbooks.

In Bahrain there are two statutory examinations: at the end of intermediate schooling (ie at the end of basic education); and at the end of secondary schooling. The current national assessment system, now in place for some ten years, was introduced for a variety of reasons: to meet changes in the structure of education in Bahrain, to employ a more diverse range of assessment measures, to grant more equity for students and to ensure greater credibility for the results. It involves the assessment of all students in the public schools in Arabic, English, Mathematics, Science and Social Studies. The final result incorporates the students' performance in an externally set and internally marked public examination, in an internal mid-term examination and as determined on the basis of continuous assessment. The external examination contains both multiple-choice and open-ended items. There are no explicit or formal steps to ensure that tasks are free from forms of bias.

(iii) **England and Wales**

At both primary and secondary level schools currently offer a largely common curriculum of ten National Curriculum subjects and Religious Education. There are some constrained options, particularly during the last two years of secondary schooling, and schools do also offer subjects which are not part of the National Curriculum (for example, Classics, Personal and Social Education and Environmental Studies). The content, but not the teaching strategies or the resources, of National Curriculum subjects is defined by government legislation. Private schools may choose to offer the National Curriculum, in whole or in part.

Schools may seek exemption from the National Curriculum requirements, either wholly or in part, for students with physical or mental disabilities or with learning difficulties. Especially able students may move through the National Curriculum requirements at an accelerated pace. Government funding is available to provide additional teaching support for students whose first language is not English.

The implementation of the National Curriculum represents a major shift in the curriculum and in the extent to which it is subject to legislation. It follows a wide range of curricular initiatives in the past 20-25 years as perceptions about the needs of students and of society have changed. Innovation has been designed, for example, to meet the needs of an increasingly pluralist society; to introduce greater commonality; to achieve better breadth and balance; and to chart continuity and progression more successfully.

A national structure for assessment exists as part of the National Curriculum. Subject assessment is based either on a 10 level scale over the eleven years of compulsory schooling; or on performance statements at the end of Key Stages (at age 7, 11, 14 and 16).

(iv) **Guatemala**

The overwhelming proportion of primary schools, public and private, follow a common curriculum - Mathematics, Natural Sciences, Social Sciences and Spanish as the basic subjects, with English and Religious Education as the most common optional subjects. The basic subjects are compulsory, but legislation allows schools some freedom to adapt the curriculum to meet the particular needs of students and of the local community. All schools, including all private schools, are required to conform to curricular legislation or to have their curricula approved.

At secondary level, students follow a largely common curriculum in general culture, with opportunities in a range of areas covering aspects of - business, industry, agriculture and animal husbandry. As in primary schools, the curriculum is determined by government legislation.

There are few specialist schools to meet the needs of pupils with mental or physical disabilities. For pupils with learning difficulties, or for gifted children, the curriculum is not modified and few schools offer specialised support. In areas where there is a large Indian population, there are a number of bilingual schools. Here the four basic subjects are taught in both languages for the first three grades of basic primary education. Elsewhere, Indian students may receive special Spanish courses.

Within the last 20-25 years, workshops to deal with practical everyday life have been added to the primary curriculum, and greater emphasis has been placed on the pupils' psycho-motor development. There has been some re-organisation of subjects to focus on learning processes and child development. At secondary level, there have been few changes beyond the introduction of some elements of computer science in some public schools.

A national assessment structure regulated by the government has been in place since 1986. The reasons for its introduction were: to satisfy the need for change in the national education system; to achieve greater uniformity into the evaluation of academic achievement; and to establish it on a more technical and scientific footing. The purpose of the assessment is to determine promotion to the next grade.

Assessments are internally set and administered within the school on the basis of regulations established by the government. Methodology guides indicate the need to assess performance in areas other than the purely cognitive and attempt to establish a more common assessment format - as, for example, the balance between written tests and observation of behaviour as a means of determining achievement. The report notes, however, that official regulations are not always followed. Comparison of results between schools or from one year to the next, therefore, is seldom possible.

For academic subjects, assessment is mainly through short answer written questions; less commonly through other forms of written or oral testing. Performance in other curricular areas is on the basis of practical work or through check list observation. No steps are taken to eliminate forms of bias.

(v) **Israel**

Education in Israeli schools is compulsory for all children between the ages of 5 and 16. Following the reform programme instituted in 1968, there has been a gradual transition from a format consisting of eight years of primary school and four years of secondary school to one of six years of primary school, three years of intermediate school and three years of secondary school. For the purposes of this report, primary education is defined as the first nine years, and secondary education as the last three years, of study.

The Israeli educational system is characterised by centralisation on the one hand, and complete freedom and liberalism on the other. There is no binding legislation over curriculum or assessment. As a result, there is a wide variety of curricula, teaching methods and assessment procedures.

Within primary education, several subjects are compulsory - Mathematics, the mother tongue (Arabic or Hebrew), English (from the fifth grade), Arabic (in the intermediate school), Bible (from the second grade), Local Studies, Social Studies (in the primary grades) and Science. For these subjects there is generally a basic core which must be taught plus optional elements. For some the priorities are clearly established; others identify a list of goals to be achieved. Such goals encompass values, attitudes and skills as well as cognitive objectives.

Within this framework schools are afforded considerable latitude to adjust the curriculum to students' particular needs and abilities. They have a good deal of autonomy on matters of curricular planning - the emphasis placed on each subject, the number and distribution of hours given to it - and of pedagogy. This autonomy is greatest in special schools - special either in the sense that they focus on a specific field of study, or that they provide for students with marked learning disabilities. Although no alternative curriculum is specified for gifted children, many schools and local bodies run classes, develop material and organise programmes designed especially for them.

The secondary curriculum includes mandatory subjects, making up some 75% of the study units in general schools and approximately 55% of study units in vocational schools, and electives. Greater flexibility is afforded by the provision for each of the mandatory and elective subjects of at least two levels of instruction and, within some subjects, of a wide variety of choices.

In recent years Israel has absorbed many thousands of immigrant children into its schools. They were enabled to come to terms with their new situation by, for example, absorption classes, additional instruction and the temporary addition of Hebrew language lessons to their programmes of study.

For the past 25 years or so, coinciding with the founding of the Curricular Division, the system has been very dynamic and there has been considerable change in the primary curriculum. New subjects (Media Studies, Health Education, Land of Israel studies) have been introduced; schools have been encouraged to put forward curriculum proposals; pedagogy has been re-appraised - in particular in an attempt to achieve a better balance between investigative and instructional approaches to teaching and learning.

Similarly, there have been a number of reforms in secondary education. These have been designed with the aims of: improving scholastic achievement; of bringing about the integration of students from varying socio-economic backgrounds; of adapting technical and vocational studies to the requirements of a modern economy; and as part of a trend towards increased pluralism and greater autonomy for individual schools. Conversely, the increase in the scope of the mandatory core, reflects the countervailing trend towards curriculum centralisation.

Although national achievement tests do exist, Israel does not have a statutory assessment system for testing student achievement. The structure of examinations and diplomas does not have a basis in law. Nevertheless, testing and evaluation is of great public concern. Opinions vary, but most schools would be likely to favour legislation in favour of national tests. From time to time the Ministry of Education issues regulations for educational assessment which, although not requiring that tests be administered, provide guidelines on their operation. These guidelines relate only to classroom-based assessment.

(vi) **Malaysia**

Within the three types of primary school - National, National Chinese and National Tamil - although the medium of instruction varies, the curriculum is common. The system of education is based on the rationale that the primary school should provide basic education. In view of this, all subjects are taken by all students, including those with particular educational needs. The subjects are: Bahasa Melayu, English, Native Language, Mathematics, Arts, Physical Education, Music, Islamic or Moral Education and (from grade 4) Man and His Environment. In the case of gifted students, schools provide enrichment activities. The curriculum in all schools, national and private, is required to be approved by the government.

At lower and upper secondary levels the core curriculum remains largely the same as in primary schools, with the addition of Science and the substitution of History and Geography for Man and His Environment.. All government assisted secondary schools use Malay as the medium of instruction, with Chinese and Tamil languages as additional subjects. Schools also offer a range of elective subjects within the areas of: humanities, vocational and technological education, science, and Islamic religious knowledge.

Elements of commerce and manipulative skills - to be taught as Living Skills - have been introduced into the primary curriculum in the last ten years. No elements have been removed.

Living Skills has also been introduced as a subject at lower secondary level; and a number of pre-vocational subjects removed from the curriculum. The reason for this was a concern that students should not specialise in specific vocational areas at too early a stage.

At upper secondary level, the curriculum has been extended by the addition of a number of elective subjects.

The Primary School Assessment Test is a national examination centrally set by the Malaysian Ministry of Education's Examinations Syndicate for the purpose of assessing the performance of all students at the end of six years of primary education. Originally its purpose was to select students for entry into secondary schools. But, as promotion into secondary education is now automatic, the examination is geared more for the purpose of some remedial work rather than for

selection. However, the results of this examination are used for selection into the residential schools specially established for rural students.

For primary students, assessment is in the basic skills - reading, writing and Mathematics - and in Science, Geography and History through the course on Man and his Environment. The singular nature of the national curriculum is perceived to warrant a centrally conducted examination with a singular instrument common to all. The only divergence from this is that the examination is set in three languages - Malay, Chinese and Tamil - depending on the medium of instruction of the students. At both primary and secondary level, forms of bias are guarded against by establishing groups of practising teachers, representing various economic, social and cultural backgrounds, who generate and moderate items at the initial setting stage. Items are also pre-tested and reviewed.

Assessment tasks are written, multiple-choice and essay response; or, in the case of languages, oral. Multiple-choice responses are marked using answer keys established during the pre-test and confirmed by post test analysis, and are marked mechanically. Students' essay responses are marked against a set of criteria using a 'highly objective' marking scheme. Reliability is also achieved through co-ordination meetings at various levels of marking. For oral and, at secondary level, practical tests neutral assessors evaluate and moderate samples of students' performance.

At secondary level there are two major national assessments - at the end of grades 9 and 11 - centrally conducted by the Examinations Syndicate. All students are involved. The purpose of the first assessment is to select students for entry into grade 10 and for placement into the arts, sciences or vocational/technical streams. The final assessment is for the purposes of certification, as an entry qualification into certain areas of employment, and for selection into higher education.

> *'Experience in Malaysia shows that there is absolute dependence on these centrally conducted examinations. Results of school conducted examinations as indicated in School Leaving Certificates rarely have any significance either for further academic advancement or job selection.'*

Assessment covers both core and elective subjects. It includes written, practical, oral and, in the case of elective subjects, course work tasks. The Examinations Syndicate lays down guidelines for the conduct and assessment of course work. But, even so, it is perceived that with the introduction of course work 'the process of decentralisation has taken a brave new step towards a degree of differentiation.'

(vii) **Namibia**

Prior to independence, the curriculum was based on the Republic of South Africa model. It is in the process of change and in state schools, at both primary and secondary level, is wholly determined by government legislation. This does not apply to private schools.

Primary education is being divided into Lower (grades 1-4) and Upper (grades 5-7). Lower Primary will offer a largely common curriculum; Upper Primary will offer a common core of subjects plus a very limited choice of optional subjects. In both Lower and Upper Primary, the common subjects are: English, a language other than English, Mathematics, Arts, Religious Education, Physical Education and Health Awareness. In addition - at Lower Primary level,

students take; Literacy, Handwriting, Environmental Studies, Craft and Technology; and at Upper Primary level: Science and Social Studies, with Craft and Technology, Home Ecology and Elementary Agriculture as options.

Secondary education is being divided into two years of junior secondary and two years of senior secondary schooling. Junior secondary education offers a common core of subjects with students' options limited to the choice of additional languages or areas of pre-vocational study. At senior secondary level students make constrained choices of five or six examination subjects from within four fields of study. Certain non-examination subjects (Physical Education, Moral and Religious Education) will also become compulsory elements of the curriculum.

Students with physical or mental disabilities receive special education. Students with learning difficulties are catered for by such means as compensatory teaching and specially adapted learning materials. There is an intention that gifted students should be given more challenging tasks in order to reach their optimum achievement level - but this remains only an ideal in most schools.

All learning at the early stages (grades 1-3) is in the mother tongue or English. Thereafter English is the medium of instruction for Mathematics, Science, Health Education, Social Studies and English; other subjects may be taught in the mother tongue or a familiar local language.

Namibia is in the process of introducing a national assessment system as part of the drive towards achieving four broad goals: to enhance equity in access to education; to achieve transparency and democratic participation; and to strive for relevance in content and quality learning outcomes. The main purpose of student assessment is seen as being to provide information by which to monitor educational policy, to determine how well strategies and plans are executed at regional and school level and to identify the impact of reform initiatives on students and the community in general.

National assessment is currently restricted to the primary level. It is to be extended to the secondary level in 1996 - 98. The initial, and now completed, stage was to provide baseline assessment data of students' performance in English, their home language and Mathematics. Science is to be added to these subjects from 1995. Testing is by means of multiple-choice items for Mathematics (and ultimately for Science); and by a combination of multiple-choice and open-ended questions for languages. Assessment is largely criterion-referenced against subject related learning objectives. Norm-referencing is used to determine gaps and differences between regions, and to identify schools in critical need of intervention.

(viii) **Poland**

Polish primary (fundamental or ground) schooling begins rather late, at the age of seven, and lasts for eight years. Some 15% of young people proceed to four years of general education in the grammar schools. Most of these students, unlike their colleagues in the secondary vocational schools, will be candidates for universities and other higher education institutions.

> *'Under socialism, for more than forty years, our primary school was uniform in respect of curriculum. There was only one common curriculum irrespective of student abilities and teacher competences. The selection of teaching content was considered a subject of the highest governmental significance and every curricular change was ratified by the top political authorities. In the early part of the period the school administration constituted a kind of curricular police investigating any discrepancy between commanded and implemented curricula ... This approach made our teachers passive and reluctant to any curricular change ... After 1989, a major revision of the uniform school model was announced ... Students' and teachers' freedom has been enlarged.'*

The main change that has come about is in the removal of 'intra subject constraints', rather than in any reduction in the range of prescribed subjects. Only two subjects - Religion (or Ethics) and Computer Science - are optional. A 'curricular minimum', occupying some 70% of the curriculum, remains. But the concept of this 'curricular minimum' is ill-defined and ambiguous. Generally it is interpreted by schools as specifying content rather than as identifying a set of learning goals.

> *'The typical Polish (primary) teacher believes curricula are ... made for him/her rather than by him/her and he/she feels authorised to small displacements and corrections only, leaving the overall structure and the volume of the total untouched.'*

Within grammar schools, students may select courses which have a humanistic, a mathematical-physical, or a biological-chemical focus. A few schools also offer courses where the focus is on computer-scientific, managerial, athletic, artistic or linguistic studies. Courses do not differ markedly in the range of subjects they contain but rather in the depth to which particular subjects are studied.

The most notable secondary curricular innovation in recent years has been the re-introduction of the study of religion as an optional subject. Other changes have been rare and only cautiously introduced. This caution is a result of the concerns of higher education institutions about the academic qualifications of would-be entrants. They would be quick to object to any perceived reduction in the demands of the secondary curriculum.

The newly established private schools, at both primary and secondary level, and specialist schools for students with physical or mental disabilities do not commonly take such a restricted view of the curriculum. Curricular adjustments are also made for gifted students and for students with learning difficulties.

There is no national or regional structure of assessment in primary and secondary schools in Poland, but there are statutory regulations to which schools are required to refer. The regulations define a six level grading scale, a descriptive grade or form of descriptive appraisal, and procedures for managing the assessment process. They do not determine the means of testing. This is the choice of the individual school or teacher. The absence of a common testing system or of specimen test items leads inevitably to a wide variety of interpretations of the grading scale.

Theoretically, the strategies used by teachers are criterion-referenced in that students' responses are assessed in terms of the curricular knowledge and understanding shown rather than against other students' scores. This principle is firmly embedded in Polish instructional theory. The school system was never influenced by differential psychology and norm-referenced measurement is entirely strange to it. In practice, however, assessment on the basis of comparing students' answers exists - especially where curricular requirements are stated only vaguely or where they are unrealistically high.

In addition, in general education secondary schools, there are two statutory multi-subject examinations. The first is an entrance examination; and the second, the Matura, is a final examination success in which is a requirement for entry into higher education and post-secondary training institutions. Both are internal school-based examinations, although parents and school board representatives may participate as observers.

(ix) Scotland

Primary education begins at about age five and continues for seven years. The official age of transfer to secondary education is 12. Compulsory secondary education lasts for four years, but a high proportion of students remain in secondary schools for a further two years beyond the statutory leaving age.

Students in primary education and in the first two years of secondary education follow a largely common curriculum, but one that is mainly determined by the schools themselves. Schools have the freedom to organise their own programmes of study using centrally produced curricular guidelines designed to provide the framework for a coherent and consistent curriculum. Guidelines set out the minimum proportion of time to be given to each curricular area and include a 20% flexibility factor of total curricular time. Each area of the curriculum is divided into broad attainment outcomes. Within each outcome are a variable number of strands on aspects of learning. Most stands have attached to them attainment targets at five levels of attainment. These targets represent progression of learning within the strand.

Responsibility for students with mental or physical disabilities, with learning difficulties, or for whom English is a second language rests with local Education Authorities. Wherever possible students with special needs are integrated into mainstream education. Some areas cater specifically for gifted students; but, in general, it is for the classroom teacher to make provision for such students.

In general terms, over the past 20-25 years, the primary curriculum has necessarily become broader to take account of the needs and development of contemporary society. This has been most notable in such areas as design and technology, social and moral development, and computer

education. At secondary level many new examination subjects have been introduced, and some few have been removed. The changes reflect swings in the popularity of subjects with students, advances in the nature of some disciplines and the re-focusing of certain areas of the curriculum towards a wider range of ability.

In Scotland, a national assessment structure for primary schools - but not for secondary schools - has recently been introduced. Its aims are to provide for more coherent and consistent assessment policies in schools and to provide parents with better information on students' progress. There are national tests in Language (reading and writing) and Mathematics. As legislation currently stands, it applies to all students in the fourth and seventh (final) years of primary schools, except for those who are formally recorded as having special educational needs.

Testing, within the context of teachers' on-going continuous assessment, is designed to confirm teachers' judgements about students' progress and development needs in relation to nationally agreed attainment targets. Tests are broadly criterion referenced with individual test items designed to test specific attainment targets. They are graded on a five point scale to reflect levels of attainment defined in the relevant curricular guides. Teachers choose tests at the level of difficulty most appropriate for each individual pupil, and also choose appropriate test contexts from the national test catalogue. In Mathematics and reading, a student must achieve a pre-determined threshold score in order to be awarded the level at which they sat the test. Students can only be awarded the level at which the task was set. In writing, different aspects (for example language and structure) are assessed according to level related criteria. These are then combined to give a single reported level.

(x) **Slovenia**

Elementary schooling (primary and lower secondary) is compulsory for children aged 7-15. Secondary education in gymnasia, professional or vocational schools, for students aged 15-19, is not compulsory - but, in practice, it is undertaken by some 90% of young people. The curriculum is wholly determined by government legislation.

Within elementary schools, students follow a given range of subjects for a specified amount of time during the school year. The subjects are in six curricular areas: languages and the arts; socio-economic studies; natural sciences and mathematics; production and technical studies; health and physical education. In addition there are a number of compulsory supervised activities.

At secondary level, curricular options - alongside a number of compulsory subjects (Slovene Language and Literature, Mathematics, first and second foreign language, History, Physical Training and Art) are introduced only in the fourth year.

The education, care and training of mentally and physically handicapped students is carried out in special development groups and in specialised institutions. Students with learning difficulties and gifted students receive additional instruction and attention within mainstream schools; and, in the case of gifted students, may participate in extra-curricular activities organised by music schools, cultural and sport associations, research camps and the organisers of various competitions. Gifted students may be given state scholarships at the end of elementary schooling.

In elementary schools there has been little change in the past 20-25 years. Nothing has been added to or removed from the curriculum other than the removal, for political reasons, of Serbo-Croation language. The weighting given to individual subjects has changed only slightly. At secondary level there have been two major structural changes: the introduction of a common curriculum in the 1980s; and the division of secondary education into general, professional and vocational schools in the 1990s. The first change was introduced on the principle that all schools should lead to the world of work as well as enabling the transition to higher education; the second reflected the concern that this was an unrealisable goal.

Students leaving elementary school and attempting entry into upper secondary schools in Slovenia take an examination in two subjects - Language and Mathematics. Assessment tasks are externally set by the Ministry of Education and consist of multiple-choice and short answer questions. It is a new system, introduced two years ago, for which students and teachers were ill-prepared. At this stage, therefore, any evaluation would rest largely on conjecture.

Within upper secondary schools assessment is at present internally based. But, as a result of anxieties about standards and concerns about the existing system's effectiveness for determining entry into higher education, a new external Matura examination is to be introduced in 1995. Currently students are examined in four subjects on the basis of both written and oral tasks.

COMMENTARY

For many countries, the last quarter of a century has seen marked changes in both the nature and the size of the school population. In all countries there is concern that, there should be a minimum level of basic eduction. In some of the countries involved in the survey this has posed, and indeed, continues to pose, considerable challenges. In Guatemala it requires the education system to come to terms with the needs of 23 ethnic groups speaking 23 different Indian languages. Namibia is attempting to overcome the educational discrepancies which exist between rural and urban areas or which are a residue of colonialism. The structure of the Israeli education system has undergone change in order to meet the needs of large-scale immigration.

Especially, although not exclusively, in the more industrialised countries, there is pressure to drive up the rate of participation in secondary schooling to a level previously found only in countries such as Japan and the USA. Pressure to expand the educational system has come partly from increased individual demand; and partly from policy decisions on the part of governments who see economic growth and educational growth as intimately linked. In Australia, for example, in 1986 less than 50% of secondary students remained in school until the twelfth grade. By 1992 the proportion had increased to 75%. The aim is for this to rise to 90% by the end of the century. In Israel the number of students in secondary education grew by 64% between 1979/80 and 1990/91.

One effect of these developments - in particular their cost and their perceived significance for national prosperity - has been to intensify the educational debate. Participants in the debate - educators, politicians, parents, industrialists - have become more searching, and potentially more critical, in the questions they ask of schools and their curriculum.

The response to questions about the form the school curriculum should take has commonly been to:

- establish a greater degree of central or governmental control. This is now so, for example, even in England and Wales where there was a long history of regarding the possibility of government involvement in the curriculum as unwarranted and potentially dangerous interference. The introduction of the National Curriculum in 1988 gave government the deciding voice not only in what should be taught but also in what should be assessed; and, to a large extent, in how it should be assessed. This is also true of other countries involved in the survey such as Bahrain, Guatemala and Malaysia. Less commonly, as to a degree in Bahrain, central prescription extends also to matters of pedagogy. Of the countries in the survey, only in Scotland does government play an advisory, rather than a regulatory, role in curricular decisions.

- attempt to define more tightly a curriculum entitlement for all students which identifies essential knowledge, understanding and skills. In part these attempts seek to ensure at least the minimum levels of achievement necessary for vocational or social participation; but in part also, as for example in Namibia, they are concerned with notions of individual equity.

- evaluate and if necessary adjust the curriculum in the light of perceived pedagogical or epistemological developments. In Israel, for instance 'originally the (Curricular) Division's conception of learning was that it could only be attained through discovery and investigation with an emphasis on cognitive skills ... Over the years, however, there developed a gap between this approach and others which held that not every child should be expected to be a 'researcher' in every subject. Subsequently, greater balance was created between these differing educational approaches'. Or, in Bahrain, 'in the 1970s the subjects in the primary school were disconnected while they became more integrated in the late 1980s and the beginning of the 1990s.'

These responses have raised two problems in particular:

1. In an OECD Report, Curriculum Reform: An Overview of Trends (OECD, Paris, 1990), Skilbeck noted that 'much of the national level debate seems to leave the schools, the teachers, students, parents and the local communities unaffected.' He warned that 'there is a real risk, if such approaches are not complemented by equal attention to the processes of change within the schools, of a widening gap between policy directions and the classroom reality.' There was evidence of this gap between central curriculum policy and classroom practice in a number of responses, most notably those from Guatemala and Poland.

> *'(In Guatemala) it is necessary and urgent that curricular adaptations should be implemented; what teachers teach the children should be adapted to their language, culture and region. This readaptation is the one envisioned by the Ministry of Education but it has not been put into practice.'*

Echoing Skilbeck, the Polish response urged that 'Adopting the principles of democratic education, we should make teachers feel entitled to goal and content selection.'

The curricular gap clearly contributes to, but is not the sole cause of, the assessment gap between central intention and school reality. In Guatemala 'what has been legislated is not done in more than 85% of the classrooms.' Some of the reasons for this, as seen in Poland, were graphically and sympathetically described.

> *'The characteristic features of the typical public schools are: (i) defensive attitude against their social environment ... and (ii) against the higher school authority ... (iii) exacting compulsory education from the students. In this situation, assessment loses its statutory function of supplying information on student cognitive achievement, which is much needed by students (and their parents) to design further learning, and it becomes mainly (i) a teacher's weapon against the social environment, (ii) his/her battle-shield against superiors, and (iii) the way of everyday balancing the teacher-student relations. It is quite easy to sit in armchair judgement of teachers abusing educational assessment but it is much harder to get rid of the abuse in an overcrowded, poorly equipped school located in an educationally unfavourable environment, especially when the teacher's subject-matter competence is not very high.'*

But, as the Guatemalan response indicated, the gap between intention and reality may occur at an earlier stage than the level of the classroom. In 1988, the government announced its intention to establish a national examination system for each cycle of the educational system to provide information about students' learning. The aim was to identify areas of curricular weakness at national, regional, local and institutional level in order to take remedial action. Some five years on the system was still to be developed.

2. There can often be considerable tension between the pull of curriculum reform and the pull of assessment. This tension is at its most acute when, for example, curriculum reform is moving in the direction of enhancing students' problem-solving or social skills, their powers of creative thinking or analytical reasoning - and assessment is seeking the certainties about student performance which the national interest is seen as demanding. Stated simply, many of the responses identified an ever-present dilemma of choosing between a worthwhile curriculum dubiously assessed, or a less relevant curriculum where the assessment is seen as being more secure.

Nisbet (1992) argued that 'because assessment is a powerful determinant of what is taught and learned, there is a need to plan assessment as an integral element in curriculum reform.' He wondered whether 'perhaps we can introduce new models of assessment which will help to implement the curriculum reforms which we see as desirable'; and saw the need for 'assessment procedures ... to be redesigned to give proper place to the skills and competences which are most relevant and valued in today's society.'

Many of the responses gave added weight to Nisbet's plea. As well as inhibiting curricular reform, some respondents also feared that existing assessment arrangements were inimical to improvements in teaching and learning strategies.

The Australian response, however, suggested that the issue there was not primarily to do with a dearth of appropriate assessment strategies, nor with teachers' lack of assessment awareness.

> *'At the secondary level, as at the primary level, some progress has been made in promoting benign, positively-oriented assessment and reporting practices which operate widely and generously across the cognitive and affective domains, to replace the more limited and limiting aspects of previous practice ... A greater number of teachers now than some decades ago have a rich repertoire of assessment strategies to put in place.'*

Although the Australian experience would find some echoes in countries such as England and Wales, it was far from universal amongst the respondents. Further, the Australian response suggested that it might be deficiencies in the curriculum which were inhibiting the deployment of better assessment - rather than the other way round.

> *'The lack is of cross-curricular opportunities for these (assessment) strategies to operate, giving students a more holistic or integrated view of learning, not necessarily lack of assessment expertise.'*

Clearly, however, achieving a proper match between curriculum and assessment is not solely a matter of designing better assessment models and strategies. It has much to do with what Nisbet refers to as a country's particular 'assessment culture'. This embraces practices which have acquired credence over time; it indicates public and professional attitudes towards ability, standards and the role of examinations; it prioritises the purposes of assessment; it reflects values and socio-economic structure. These are issues which are considered in the next section of the report.

CHAPTER II

ASSESSMENT: POLICY, PRACTICE AND PURPOSE

Section 1 of the questionnaire asked for information about the nature of the school curriculum at primary and secondary level and the associated assessment arrangements in the participating countries. This has been summarised and commented upon in the preceding chapter. Section 2 invited comment from the respondents upon the following four main issues relating to current assessment policies and practices within their respective countries:

i) The nature and usage of conservative, mainstream and radical assessment practices within both the primary and secondary sectors.
ii) The degree to which a gap existed between the stated purposes of assessment and the uses actually made of assessment results in practice.
iii) The extent to which existing assessment arrangements met the needs of contemporary society and prepared young people for life as adults.
iv) The changes that were needed to current assessment arrangements and what pressures or support existed for and against change.

In relation to all four issues the respondents were encouraged to provide reasons and to express their own views and those of the people they had consulted. They were thus asked both to describe and to evaluate and in order to do this they had to define the terms used. What, for example, did 'mainstream' mean in practical terms and did it/should it mean the same in Guatemala as in Australia or in Bahrain as in Poland - all countries with very different cultures, history, resources, geography and population?

Given this situation, one might have expected to see significant differences in the answers provided such as would make a coherent overview difficult. In the event, the similarities in a number of key respects which interlink and overlap were much more marked than the differences. These were as follows:

1. The range of purposes for which the results of assessment were being used and the relative importance attached to these purposes.

2. A relatively small degree of mis-match between the stated and the actual uses to which the results of assessment were being put.

3. The use of a relatively limited range of assessment modes both in the classroom and in large scale external programmes. This limited range moreover operated within an equally limited range of curriculum content, particularly at the primary level.

4. A general concern that the current assessment arrangements did not meet the future needs of young people particularly well.

5. A marked unwillingness to support changes within assessment, particularly by teachers.

1. PURPOSES

All the individual country studies referred to three main purposes to which the results of school assessment were put, although they did not always use the same terminology. These were:

selection, which often involved promotion within as well as between institutions;

monitoring standards in various ways in respect to individuals, to institutions and in a growing number of cases to national education systems; and

improvement in the quality of learning.

By far the most common use of assessment was for selection, although the point of impact varied considerably. For example, in a country like Guatemala where only 31% of the relevant age group go on from the primary to the secondary basic cycle of education and only 12% to the secondary diversified cycle, selection and hence assessment at the point of transfer between secondary and primary becomes absolutely crucial for a child's future. A similar, but less acute situation, obtains in Namibia and in most developing countries where resources simply do not permit the building of enough secondary schools and where, on occasion, there is also limited primary provision. In countries like Australia, Israel, England, Wales and Scotland, where an increasingly large percentage are staying on voluntarily into post-compulsory education, the selection/assessment impact is primarily felt at the secondary tertiary dividing line. In other countries, like Bahrain and Malaysia, selection remains important at more than one point within the system. In some countries, as for example in Guatemala, Poland and Namibia, promotion from year to year within institutions is also of critical importance.

The second most common purpose was that of monitoring - increasingly at the national level as governments of all political hues become more interested in ensuring that the education system is providing a good return for the tax payers' money and that education is serving the national interest (variously and often vaguely defined). In some countries, monitoring has placed considerable emphasis upon monitoring the teacher and has, therefore, been associated with teacher appraisal. This, however, was not an issue in the reports in this study, with the exception of that on England and Wales. The growth both of international comparative studies like those conducted by IEA in which a number of the countries in the sample are taking part, and of national assessment programmes like the National Assessment of Educational Progress (NAEP) in the USA and the work of the Assessment of Performance Unit (APU) in England and Wales, although both very different, have been a factor in the increasing use of assessment for monitoring. It has certainly influenced thinking and given it a sharper edge in countries like Israel. The main reason, however, in the countries covered in this study for the increased importance given to assessment for the purpose of monitoring, has been the introduction of mandated national curricula. These have been introduced either initially from scratch, as in Bahrain; or as a major policy initiative as in England and Wales; or in the form of greater regulation of existing common provision, as in the case with almost all the other countries covered in this study, particularly at the secondary level. (See the first chapter for details).

Assessment for the improvement of the quality of learning, particularly in any systematic way, was referred to far less frequently although there was more reference to it at primary level and there was a notable exception in terms of emphasis in the Australian report. This apparent neglect contrasts sharply with the growing shift referred to in many of the reports away from norm referencing towards assessment models and descriptions based upon student outcomes although the changes were not often clearly defined. There has also been increased interest, notably in England and Wales, but also in other countries such as Australia and Malaysia in records of achievement or profiles and portfolios both of which lay stress upon descriptive assessment and a greater range of evidence and have, in consequence, considerable potential for involving individuals in their own assessment.

2. MIS-MATCH BETWEEN PURPOSE AND PRACTICE

There is little evidence in the reports themselves that there is a significant mis-match between the stated and the actual uses of assessment results - with one exception. Most reports, particularly in respect of primary education, referred to diagnostic or formative uses, but it is clear that the methods of assessment, the large scale use of marks and grades and the ways in which the results were recorded did not lend themselves to any systematic use being made of assessment information to improve the quality of either teaching or learning. Reference to this use of assessment is also conspicuous by its absence in relation to teacher training.

The Israeli report provided an interesting description of what came to be called 'feedback tests'. This illustrates the problems associated with assessment designed to achieve diagnosis through central arrangements and in which there is some confusion of purpose as to whether the diagnosis is primarily concerned with individual students or with institutions.

In 1991 the Ministry of Education introduced assessment arrangements to provide feedback on the extent to which the educational system was 'fulfilling its public obligations' with respect to academic achievement - in particular arithmetic and reading comprehension. The arrangements are designed to provide comparative data between districts, cities, regions and schools, hence reflecting the accountability of the educational system as a whole - a process most of the respondents to this study would describe as monitoring. They have a number of interesting and relatively unusual features:

i) Involvement in the scheme is not required by law but all the schools within the state sector, to which it is confined, take part - nobody ever seems to pass by assessment if they have the chance!

ii) The scheme was to be independent of, and not financed by, the Ministry of Education. In practice, however, the Ministry provides partial funding and services such as computing and statistical support. This raises the point of whether any scheme designed to address issues of national accountability can exclude the government.

iii) <u>All</u> students, both Jewish and Arab, within the state sector are examined at a particular grade level determined annually. What is tested might also change on an annual basis, but this has not happened as yet. In 1991 all students at grades 4 and 5 were tested on arithmetic and reading comprehension. In 1992 all grade 3 students were tested in these

subjects. In future it is intended to administer tests in the Intermediate grades, as well as in other subjects - possibly Geography and the Natural Sciences. This raises the issue of what the implications might be for classroom practice of testing only a limited range of the curriculum.

iv) The tests, which consist of multiple choice items only, are in two sections: the first is a minimum criterion referenced test, the second is a norm based test. Each student, therefore, receives two grades: a pass/fail grade on the criterion referenced section and a second grade based on the percentage of correct responses compared with the norm. This raises two questions: how appropriate for purposes of diagnosis is the mode of assessment used; and how should information which is to be used for diagnosis be presented?

v) The principal objective is to see whether lessons learnt from the previous year's tests are put into practice in subsequent years, and it is for this reason that the grade levels change annually. The second objective is to use the scores obtained by students as criteria for allocating resources; so far this has not happened. Results are reported to parents, students and schools. This raises the question of what assistance schools need in the diagnostic use of tests.

Almost immediately following the introduction of the scheme a number of schools decided to develop their own additional tests in the two chosen curricular areas. These increasingly have come to be used as preparation for the 'feedback' tests themselves and have led to a substantial investment of time and other resources. The 'feedback' tests are limited in nature, consisting as they do entirely of multiple choice items, and this restricts the range of feedback that is possible. The school averages and the percentages of those who have passed and failed commonly appear in the press. This results in a great deal of ill-informed criticism and comment and consequential resentment amongst teachers who see themselves as being monitored by the tests. Thus 'Low Stakes' assessment designed to provide useful diagnostic feedback was in a very short space of time turned into a 'High Stakes' programme in which all parties feel threatened.

It could be argued that the original purpose of 'diagnosis' had been subverted. But it is equally arguable that the original purpose itself lacked clarity and failed to recognise that a distinction must be made between diagnosis of the system (monitoring), which is the prime and proper concern of the state; and diagnosis of the strengths and weaknesses of individual students, which is the prime and proper concern of the teacher.

3. MODES OF ASSESSMENT

The Malaysian response to the questions on modes of assessment was typical of the majority of the responses on this topic. The main mode in formal assessment is written with some and, in a number of cases, a growing use of orals and practicals in appropriate subject areas; although countries like Poland and Slovenia have also had a very long tradition of oral assessment, particularly, at the upper secondary level.

Within written questions, multiple choice predominates, particularly at primary and lower secondary where the numbers are large. At the upper secondary the most frequent form of written question is the essay response - a mode of writing virtually unique to answering examination questions. With the exception of England and Wales, there was no reference to the structuring of written questions in an attempt to provide a greater range of opportunities for a wider range of students.

All the reports referred to an increasing focus on the affective and the psychomotor domains and to greater stress upon the assessment of higher order skills. Most of the respondents also referred to more attention being given to the specification of criteria in terms of both detail and clarity. As a result there appears to be a growing emphasis upon what students can do (outcomes) as the basis for awarding marks and grades - although relatively little detail was supplied of actual practice. Nevertheless, the overwhelming majority of the assessment described remains normative in nature. There was surprisingly little reference overall to criterion referencing - with some notable exceptions as, for example, Israel. Less surprisingly, there was little mention of self referencing except in relation to profiling in Australia, England, Wales and Scotland.

Another point brought out in the Malaysian response on formal assessment, which was also referred to by a number of other countries, was the introduction of an element of teacher assessment within formal external examinations. This can and does take a number of forms. In Guatemala, for example, the tests themselves are set by panels of teachers elected by the school faculty under the supervision of the Principal whose terms of appointment include a requirement to preside over the evaluation committee which sets the tests. In practice, however, the teachers set their own. In Malaysia and in England and Wales there are teacher assessed components (often called course work) which have different weightings in different subjects. This is also the case in Scotland where at least one of the three assessable elements in the Standard Grade is based on the whole course and marked by the teacher. In addition in Scotland, schools provide the Scottish Examination Board with an estimated grade for every student taking an examination. These estimated grades are used in the event of any disagreement. In Israel the upper secondary Matriculation examination involves a school based grade as well as an external grade. The former must take into account the average grades given by the teacher over the two years preceding the final examination. In Poland, in both the secondary entrance and the final leaving examination, the written essays are set by the teachers in subjects chosen by the students, and are marked using pooled teacher judgement. In Slovenia, within the upper secondary schools, where students are working for university studies, there are entirely internal final examinations in which the students' work is set and marked by the teachers. The results are reported by the school to individual students who receive certificates. Concerns, about lack of standards, however, have led to the introduction of a new external Matura examination at the end of secondary schooling in 1995. This will be the responsibility of a newly established national examinations centre in Slovenia.

In all these approaches, the key common factors are:

- the relative weighting given to the teacher/school contribution by comparison with the external examination contribution - and this, of course, varies;
- the extent to which a substantial component of teacher assessment reduces or enhances public confidence in the overall result (usually the latter);
- the extent to which teachers see their involvement as a worthwhile professional activity or simply as extra work (very often the latter).

Under classroom assessment, the Malaysian response made brief reference to the use of standardised tests and observation as well as to orals and practicals; but overall the written mode and formal testing, marking and recording on a regular basis predominate here also. The deficiencies of this approach are well set out in the regulations issued in 1988 by the then Director General of Education in Israel. He warned against excessive use of formal testing in classrooms and stressed the need for a much less directive and more systematic approach to monitoring pupils ongoing performance on a continuous basis.

Only in the Australian response and to a lesser degree in those from Scotland and England and Wales were there references to informal observation, flexible timescales for assessment, student involvement in their own assessment, working in groups, open ended assignments and the use of descriptive and positive approaches for the recording of student achievement. The Australian response was also the only one to stress the importance of an overall strategy for assessment, whether in countries or in schools. This is in marked contrast to the piece-meal nature of most current practices as described in all the other countries involved in the study, despite a marked and increasing degree of central involvement in assessment.

All these references emphasise the point that, rather than focus on the modes of assessment themselves, it would probably be better to consider the overall environment within which assessment is conducted, the methods used to record information about students and the strategies adopted to make constructive use of such information. It is these factors, rather than the modes, that restrict or enhance assessment provision.

Indeed, there are in reality only a limited number of modes of assessment which can be used. How they are used is what matters. An individual can be asked to do something either in writing or orally or pictorially or through first hand or simulated demonstration. He or she can respond in writing, orally, pictorially or by demonstration - all common modes of assessment in wide general use and referred to frequently in these responses. There is, therefore, no new mode of assessment awaiting around the corner to be discovered. What creates the difference is the formality or otherwise of the process; the narrowness or flexibility of the timeframe involved; the range of evidence; and the ways in which students' work is evaluated, used and reported upon. Significantly, it is in these respects that the practice referred to in the countries involved in the study is generally narrow. Moreover, it is narrow in similar ways across a group of countries which are very different. Furthermore, this shared narrow assessment framework operates in an equally narrow curriculum framework. As the last chapter makes clear, there is surprisingly little difference in the overall content of the curriculum in the ten countries involved in the study - particularly at the primary level; although there are obvious differences within subjects like History which form a part of that curriculum.

4. THE FUTURE NEEDS OF YOUNG PEOPLE

The responses commonly expressed concern that the current curriculum and assessment arrangements did not adequately meet the needs of young people today, particularly in terms of enabling them to function effectively in a rapidly changing world. A substantial number of responses implied, by omission as much as anything else, that there had been very little thinking about what changes would be required to make these arrangements more appropriate.

There are, of course, a number of aspects to this question. One relates to the kind of flexible skills which the modern world increasingly requires, others relate to adequate provision to meet cultural diversity, gender discrimination and the needs of the disadvantaged and gifted. There is moreover a distinction to be made between recognising that issues such as these are important whether considered in a social, economic, moral, or practical sense, and doing something about them at the point of assessment. It is in the latter respect that the responses suggested that current provision is inadequate and on occasion seems not to exist at all.

In its response to the questions on the appropriateness of the existing assessment arrangements to the needs of contemporary society, Malaysia stressed the importance of reasoning, thinking and adaptive skills - a view which the other respondents explicitly or implicitly supported. The problem, however, lies not in agreeing at a general level but in specifically translating such skills into curriculum and assessment provision that can be delivered in a variety of contexts. There is little or no evidence from the responses that the necessary thinking and work is currently in hand to do this.

Those countries where there are some possibilities for change at the present time either continue to polarise an artificial distinction between the academic and the vocational, as in England and Wales; or assume, as is the case with some Australian states, that education becomes increasingly a matter of vocational training as a student grows older and that the needs of vocational education ought to determine larger and larger parts of the upper secondary curriculum and its assessment. Issues such as these also lie behind the current debate in Israel over the functions or purposes of assessment. They raise the question of whether it is possible to provide a curriculum and assessment solution which will cover all requirements, or is it necessary to go for a variety of provision - with all the problems that poses with regard to esteem and reliability? What, however, is clear from the responses, is that the needs of certain key groups within society are not adequately being met at the present time, particularly in respect to appropriate assessment provision.

None of the countries in the study is completely homogeneous in cultural or ethnic terms and for a number, this is a very significant issue. For example; Guatemala with its 23 Indian tribes each with its own language; Malaysia with its substantial Chinese and Tamil communities; and Israel with its central European immigrants and Arab and Druze population which constitutes 17% of all Israeli students.

Where action is being taken, the most common assessment provision includes:

a) the availability of alternative syllabuses in areas like Religious Education;

b) translating papers into more than one language;

c) ensuring that examinations, when under construction, are scrutinised by representatives of the major groups which will be taking them.

Of these, only the last (and that only to a degree) is concerned with provision at the point of assessment.

In addition, Israel makes provision for:

d) different but equivalent examinations;

e) the granting of credit (10 marks out of a total of 100) for those answering an examination in Hebrew whose first language is not Hebrew.

Of these again, only the first makes provision at the point of assessment.

The only country in the study to provide a specific check list for analysing cultural and racial bias in test material, which research has shown to be widespread, is Scotland ironically one of the most homogeneous countries in the sample. There is also a set of more general regulations which apply to England and Wales in respect of National Curriculum assessment, and the need to take appropriate action in these matters has again been emphasised in a recent review report.

The check list used in Scotland, which is reproduced below, is an example of what is needed at the point of assessment. Experience in its use will almost certainly lead to future revision. On the evidence of the responses, the majority of formal statutory and classroom assessment referred to in this study would not meet the conditions specified here.

Checklist for Analysing Cultural/Racial Bias in Test Material

Every possible effort must be made to ensure that the units are free of cultural and racist bias. Vetters must bear in mind the linguistic and cultural diversity of modern Scottish society and the value to all children of including in the units, material which reflects such diversity. Many children in multi-racial Scotland are familiar with more than one culture.

You should be able to answer 'yes' to these questions ...

1. *Are people from a variety of cultural backgrounds equally likely to appear in pictures, illustrations and examples?*

2. *Is the portrayal of other cultures balanced and not represented as 'poor', 'primitive' or 'exotic'?*

3. *Are people from a variety of cultural backgrounds shown to be taking an equal and active part as parents and members of families?*

4. *Are people from a variety of cultural backgrounds shown to be participating equally in physical and practical activities? In scientific and technical activities? In artistic and domestic activities?*

5. *Are people from a variety of cultural backgrounds portrayed as being equally competent in both intellectual and practical activities?*

6. *Are people from a variety of cultural backgrounds portrayed as having equal status at work and at home?*

7. *Are the situations and contexts used within the experience of, and of interest to, children from a variety of cultural backgrounds?*

...and 'no' to this question.

1. *Do the texts and related activities make assumptions about cultural, class and religious backgrounds which will make them inaccessible to some pupils?*

The responses indicated that the situation in respect to specific provision to address gender bias, again at the point of assessment, is even less satisfactory. The cultural diversity of the participating countries results in a wide variety of views and practices regarding the role of women in society and in employment. Much research has shown that there are major problems in both assessment and curriculum provision in areas for example, like science and technology. Yet gender bias was mentioned in only one response (England and Wales). Again, only Scotland provides a checklist for analysing gender bias in test materials. This is again provided below.

Checklist for Analysing Gender Bias in Test Materials

You should be able to answer 'yes' to these questions...

1. *Are girls and boys, men and women, equally likely to appear in the texts, illustrations and examples?*

2. *Are females and males shown taking an equal and active part as parents and members of families?*

3. *Are females and males shown participating equally in physical and practical activities? In scientific and technical activities? In artistic and domestic activities?*

4. *Are females and males portrayed as being equally competent in both intellectual and practical activities?*

5. *Are males and females portrayed as having equal status at work and at home?*

6. *Are the situations and contexts used equally within the experience of, and of interest to, girls and boys?*

... and 'no' to these.

1. *Is the male pronoun (man, he) used to refer to all people?*

2. *Are men and women shown only in stereotyped roles (eg women as housewives, carers, secretaries and men as managers, scientists, engineers)?*

3. *Are females portrayed in more passive roles (sitting, watching) and males in more active roles?*

4. *Do females tend to be in sensitive roles and males in aggressive roles?*

Again, on the evidence of the responses most of the assessment described would fail to meet the requirements of this checklist. It would be interesting to see whether the points made or not made in these responses in respect to gender bias in assessment, would have been different if countries such as the United States of America, the Soviet Union or Sweden had been included. Particularly in the USA, there have been a substantial number of court cases in recent years relating to bias in assessment; not least because the requirements in respect to such bias have been set out clearly in writing for those engaged in test construction.

As far as providing for students with special needs is concerned, the situation is much more complex. This is for two reasons:

i) Differences in how the term 'special needs' itself is defined. This is further complicated by the greater understanding that is emerging almost daily through improvements in our knowledge of the human body and mind.

ii) The various arrangements made by different countries either to integrate students with special needs into main stream provision or to make special provision to educate them with similar students. This again, varies according to the nature of the needs and the possible effects upon other children.

Both these factors are, of course, heavily influenced by resource considerations which allow or preclude certain provision whatever may be the views and wishes of the country concerned; and by the fact that these views and wishes will themselves be culturally influenced.

The result is a wide variety of practice ranging from an apparent absence of any specific arrangements, through such things as exemption from specific activities to the provision of a special curriculum within designated institutions. This last is particularly common for deaf and blind children. Again the distinction needs to be made between addressing the issues as a social matter and addressing them at the point of assessment. And, as with cultural and gender differences, there was little evidence in the responses of any specific assessment provision for special needs - although the topic is much more widely referred to in relation to teaching and learning. Guatemala, for example, referred in its response to curriculum adaptation, although not to assessment adaptation; Bahrain referred to remedial provision in relation to reading at the primary level, but not to its assessment; Namibia identified compensatory teaching and specially adapted learning materials, but did not indicate how these are evaluated and how the students are assessed.

There was no specific reference to assessment provision for gifted students within mainstream formal external assessment. In some systems gifted students are allowed to take specific examinations early or to accelerate through the grade structure. Here it is reasonable to assume, although there was no direct evidence, that teachers would generally provide more demanding assessment in the classroom. Subject Olympics, sometimes organised nationally and sometimes by professional associations and/or commercial companies, are a feature in Slovenia and Poland. These, inevitably, attract able students and great prestige is attached to them. It is likely, therefore, that the questions asked would be demanding ones designed to identify and differentiate between able students. The same is also true of scholarship and entrance examinations for high status institutions which select on academic excellence in countries like England and Australia.

5. ATTITUDES TO CHANGE

Without exception, the responses revealed, a marked unwillingness to support change in existing assessment arrangements. This was particularly so at the secondary level, not only only amongst teachers but also amongst parents, students and the community at large. A desire to maintain the status quo was clearly identified. The reasons for this are not as straight forward as might appear at first sight and will be the subject of discussion in the next and concluding part of this chapter. The Malaysian response, however, summed up certain of these reasons very well. It noted that

> *'in the Malaysian context there is apparently no real pressure from within or without for changes. People are generally satisfied as the certificates issued by either the Examination Syndicate or Examinations Council (in the case of the HSC) have always been acceptable entry qualifications in most parts of the world. The favourable rates of success of students seeking education overseas, based on locally issued certificates, speaks volumes about the standards of education in Malaysia'.*

CONCLUSION

The picture provided of current assessment provision in the ten participating countries is not a particularly encouraging one. Moreover, it apparently obtains across a range of countries which differ very markedly in respect of such things as culture, resources and population. It is a picture which shows that assessment results are much more widely used for selection than they are for diagnosis - a situation that is widely accepted; that most of the assessment in current use is restricted in scope and operates within a relatively restricted curriculum framework; that the provision at the point of assessment is not adequately equipping young people going into the world after school and is not adequately addressing the problems of cultural diversity, gender bias, special needs or gifted students. Finally, it shows that there is remarkably little pressure for change. It is not, however, enough simply to make these points which, it must be stressed again, are based on this particular study. The points may be replicable, but the study of itself will not provide the evidence to say so with any degree of confidence. It is important, therefore, to provide some explanation for this picture and this is what the concluding section will try to do.

Hargreaves writing in 1989 stated that 'it would be no exaggeration to say that the 1980's had been the era of assessment led education reform'. This contention has been widely supported in the literature, see for example, Broadfoot (1986) and Withers and McCurry (1990); as well as in practice, see for example, the National Curriculum in the United Kingdom and some of the initiatives in the Australian states. Assessment practices and procedures have a profound effect on what goes on in the classroom - 'what you test is what you get' - and it is, therefore, not surprising that assessment has been seen as a vehicle, if not the vehicle, to promote and sustain curriculum change. A second feature of the past decade in all parts of the world has been the increased involvement of teachers in the assessment of their own students as a part, and a part that counts, in influential external public examinations.

The conservative view of assessment practice provided by the responses seems at first sight at odds with these initiatives. Kevin Piper and Barry McGaw of the Australian Council for Educational Research in a discussion paper entitled Assessment Led Education Reform; Aspirations and Limitations, which was prepared for a UNESCO African Round Table on Educational Assessment for Educational Policy and Strategy held in Nairobi in May 1991, pertinently pointed out that assessment procedures have an extraordinary capacity to linger long after their educational relevance has passed. There are, of course, numerous examples of examinations like the British General Certificate of Education at Advanced Level (A Level) which have overstayed their welcome. In consequence it is changes in assessment, rather than changes in curriculum, that are likely to attract opposition.

Amongst the examples quoted in their paper is that of Victoria, the second largest of the Australian states. Over the past five years, Victoria has been engaged in a radical and difficult reform of upper secondary schooling, the impetus for which has arisen from a very large increase in those staying on voluntarily in senior secondary schools - the 1980's saw this number double. This has meant that senior secondary education can no longer be viewed as being for a minority of those wishing to go on to higher education. The development of a more inclusive curriculum to meet the needs of this greater range of students has had to be complemented by a wide ranging reform of the assessment through the introduction of a new Victorian Certificate of Education. The curriculum reforms which were at the heart of the matter have not on the whole sparked off a great deal of controversy or debate. On the other hand, the Common Assessment Tasks

(CATS) developed to support these reforms, some of which are assessed within schools, have come under a concerted attack led mainly by the Universities. This has forced a number of concessions and modifications to the assessment procedures which will inevitably weaken the curriculum reform itself.

Analysis of the responses showed a remarkable similarity in the curriculum provision of each of the ten countries. Moreover that curriculum provision was itself fairly narrow It is perhaps hardly surprising, therefore, that the assessment model in use is also narrow and relatively static. The report abounds with references to the conservatism of teachers in relation to assessment, and also in relation to teaching and learning styles. The didactic, or 'if I am not teaching, you are not learning', approach is very deep rooted and is particularly at risk if major changes take place in assessment. Sometimes, as in Guatemala, this conservatism is supported by arrangements which guarantee tenure for life at a fairly early stage in a teacher's career. But probably the most important factor is the absence of any continuous long-term programme of staff development in assessment which links pre and in-service training. Such training as is provided rarely goes beyond assessment construction and is often designed to provide examiners for the central assessment system. Often the training itself is criticised by the teachers, as in the Australian report, and only in the Polish contribution is there any discussion about theoretical issues and the role that these might play in professional development.

The main factor, however, in the desire to maintain the status quo in assessment, appears to be the purposes to which the results of assessment are put. As already indicated, the most common use of assessment found in this study, was for selection purposes followed by monitoring with diagnosis a poor third.

This same pattern was reflected, in a conference entitled Learning, Selection and Monitoring; Resolving the Roles of Assessment run by the International Centre for Research in Assessment (ICRA) in July 1993. Of the eight countries which submitted papers moreover, only one, England and Wales, was involved in this study. Both selection and monitoring are 'High Stakes' assessment as far as the individual and indeed as far as the institution, where this is relevant, is concerned. This in consequence has major implications for secondary education in particular, as the Malaysian comment made earlier shows; notably in the lack of variety in the assessment methods used combined with an unwillingness to make much in the way of change. Those methods which are used tend in consequence to have a very high status and exercise an undue influence upon classroom practice which teachers in the main lack the will or the expertise to resist. 'High Stakes' assessment also emphasises reliability and, therefore, tends to use methods like multiple choice that maximise reliability. The desire for simple, reliable assessment is further reinforced by concerns about equity and about security, both crucial to the success of 'High Stakes' selection assessment.

The result of all this is to lower substantially the risk taking propensity of all parties involved in the assessment process and to stress simplicity and technical reliability. Neither of these aid diagnostic use which requires a much more flexible approach in both timing and method. Piper and McGaw in their paper draw attention in particular to three dilemmas which require accommodation if assessment is to contribute constructively to curriculum change and improve the quality of learning; namely, Achievement or Learning? Accuracy or Adequacy? Inclusive or Exclusive? These dilemmas, and they are very real ones, are at the heart of the Israeli debate and are present implicitly in many of the points made by the other respondents. Failure to resolve

these dilemmas would matter less if the numbers involved in examinations remained relatively stable and by implication the social groupings taking them also remained stable. This as the responses show is not the case. Change will continue to accelerate. More and more people will take and wish to take 'High Stakes' examinations and there will be growing pressure, therefore, to maintain the reliability and the neutrality - and hence, the standards. The higher the stakes, the greater will be the pressure. The result, as this study clearly shows, is and will continue to be an assessment model that is less and less appropriate for most of those taking it. One of the great ironies of assessment, as this study also underlines, is that where the need is greatest for progressive innovative curricula, so the pressure is also greatest for recognisable and comparable forms of assessment. There is no easy or quick solution to this problem. But it must lie in part in the provision of more systematic and comprehensive training for teachers in the use of assessment. Unless this occurs it will not be possible to raise the status of diagnostic assessment and without this, designing an assessment model which has the improvement of quality in respect to both teaching and learning at its heart will be impossible. The ultimate irony, of course, is that success here would lower the stakes in relation to both selection and monitoring and make both processes easier. The reverse is not, however, the case.

REFERENCES AND BIBLIOGRAPHY

Berlak, A & Berlak, H (1981) Dilemmas of schooling: Teaching and social change. London: Methuen

Black, H D; Dockrell, W B (1984) Criterion-referenced Assessment in the Classroom. Edinburgh: Scottish Council for Research in Education.

Broadfoot, P (1986) Profiles and records of achievement: A review of issues and practice. London: Holt, Reinhart & Winston.

Broadfoot, P (1987) Introducing Profiling. London: Macmillan

Brooke, N & Oxenham, J (1984). The influence of certification and selection on teaching and learning. In J Oxenham (ed) Education versus qualifications? London: George Allen & Unwin

Brown, M (1990) Graded Assessment Projects: Similarities and Differences. In: R Murphy and R Moon, Developments in Learning and Assessment. Dunton Green, Hodder & Stoughton (for the Open University).

Burkhardt, H; Fraser, R & Ridgway, J (1989). The dynamics of curriculum change. In J Malone, H Burkhardt, and C Keitel (eds) The mathematics curriculum towards the year 2000: Content, technology, teachers, dynamics. Perth, WA: Curtin University of Technology.

Claxton, C & Murrell, P (1987) Learning styles: Implications for improving educational practices. ASHE-ERIC Higher Education Report No 4. Washington, DC: Association for the Study of Higher Education.

Crooks, T J (1988) The Impact of Classroom Evaluation Practices on Students. Review of Educational Research, Vol. 58 pp 438-81.

Davies, P (1990) Differentiation in the Classroom and in the Examination Room: Achieving the Impossible? Cardiff: Welsh Joint Education Committee.

Flecknoe, M (1983) A Philosophy of Assessment. In R Moon (Ed) Comprehensive Schools: Challenge and Change. Windsor, Berkshire: NFER-Nelson.

Frith, D S and Macintosh, H G (1991) A Teachers Guide to Assessment. Cheltenham: Stanley Thornes.

Garforth, D and Macintosh, H G (1986) Profiling a User's Manual. Cheltenham: Stanley Thornes.

Gasking, D A T (1948) Examinations and the Aims of Education. Carlton, Victoria, Melbourne University Press.

Glaser, R (1990) Towards New Models for Assessment. International Journal of Educational Research, Vol 14 pp 475-83.

Goldstein, H (1993) Interpreting International Comparisons of Student Achievement. Paper Prepared for UNESCO.

Hamilton, J S; McLone, R R (1989) Education Validity and the Setting of Reliable Standards. Studies in Educational Evaluation, Vol 15 pp 295-308.

Hargreaves, A (1989) Curriculum and assessment reform. Milton Keynes: Open University Press.

Izard, J (1992) Assessing Learning Achievement. Educational Studies and Documents No 60 UNESCO, Paris.

Izard, J F - in press. Issues in the Assessment of Non-objective and Objective Examination Tasks. In: A J M Luijten, Issues in Public Examinations. Utrecht, Lemma B V.

Keeves, J P (1975) Curriculum Factors Influencing School Learning. (Paper presented at the Australian Association for Research in Education Conference, Adelaide, South Australia).

Macintosh, H G (ed) (1978) Techniques and Problems of Assessment. London: Edward Arnold.

Madaus, G F (1988). The Influence of Testing on the Curriculum. In: L N Tanner (ed.), Critical Issues in Curriculum, Part I, pp. 83-121. Chicago, National Society for the Study of Education. Yearbook of the National Society for the Study of Education, 88.

Nisbet, J (Ed) (1992) Assessment and Curriculum Reform, OECD, Paris.

Omar Hashim (1990) The Balance between External and Internal Assessments. (Paper presented at the Sixteenth International Conference, International Association for Educational Assessment, Maastricht, 18-22 June 1990).

Palincsar, A S; Winn, J (1990) Assessment Models Focussed on New Conceptions of Achievement and Reasoning. International Journal of Educational Research, Vol. 14, pp 411-13.

Piper, K (1986) Issues in evaluation and assessment: A discussion paper. Western Samoan Junior Secondary Curriculum Project (mimeo).

Piper, K (forthcoming) National curriculum: Threat or promise? Australian Education Review series. Hawthorn, Victoria: ACER.

Piper K and McGaw B (1991) Assessment-led Education Reform: Aspirations and Limitations Paper presented at African Round Table on Educational Assessment for Educational Policy and Strategy. Nairobi.

Ross, K N Mahlck, L (eds.) (1990) Planning the Quality of Education: The Collection and Use of Data for Informed Decision-making. Paris/Oxford, UNESCO/Pergamon.

School Examinations and Assessment Council (SEAC) (1990) Specification for the Development of Standard Assessment Tasks in the Core Subjects for Pupils at the End of the First Key Stage of the National Curriculum. London: SEAC.

Skilbeck, M (1990) Curriculum Reform: An Overview of Trends. Paris, OECD,

Smith, L M; Klass, C S (1989) Assessing the Whole Child. In: H Simons and J Elliott, Rethinking Appraisal and Assessment. Chapter 12. Milton Keynes: Open University Press.

Somerset, A (1993) Examinations and the Quality of Education. Draft paper written for World Bank.

Suggett, D (1985) Guidelines for descriptive assessment. Revised edition. Melbourne: Victorian Institute of Secondary Education.

Task Group on Assessment and Testing (1987) National Curriculum: Task Group on Assessment and Testing - A Report. London: Department of Education and Science/The Welsh Office.

Wiggins, G (1989) A True Test: Toward More Authentic and Equitable Assessment. Phi Delta Kappan, Vol. 70, pp 703-13.

Withers, G; Batten, M (1990) Defining Types of Assessment. In: B Low and G Withers (eds), Developments in School and Public Assessment. Hawthorn, Victoria: ACER (Australian Education Review, No 31).

Withers, G & McCurry, D. (1990) Student participation in assessment in a co-operative climate. In B Low and G Withers (eds) Developments in School and Public Assessment. Hawthorn, Victoria: ACER (Australian Education Review, No 31).

Wolf, A (1993) Assessment as Part of a System: Incentives and Control. Paper presented at International Centre for Research on Assessment Conference on Learning, Selection and Monitoring: Resolving the Roles of Assessment. London.

Wood, R (1989) Survey of Research into Assessment and Testing: A Report to the University of Cambridge Local Examinations Syndicate by PRD Ltd. St Albans, Psychometric Research & Development Ltd.

APPENDIX A

COUNTRY	ORGANIZATION	PERSON(S) RESPONSIBLE FOR COMPLETING THE QUESTIONNAIRE	OTHERS CONSULTED
AUSTRALIA	Australian Council for Educational Research (ACER)	Mr Graeme Withers Senior Research Fellow	Professional Staff at ACER
BAHRAIN	Educational Research and Development Center, Ministry of Education, State of Bahrain	Dr Nakhleh Wehbeh	Ahmed Sidawi Adviser to the Minister of Education
ENGLAND AND WALES	Mathematics Centre, West Sussex Institute of Higher Education	Mr Henry Macintosh Research Fellow	Frank Stoner Curriculum Adviser Department of Employment
GUATEMALA	Universidad del Valle de Guatemala, Institute of Research Center for Studies in Education	Dr Robert B McVean Educational Psychology Dr Otto E Gilbert Psychology Blanca de Mendizabel Licentiate Psychology	<u>Primary</u> Lissette Mejia Maria Diaz Maria de Molina Beatriz de Zelaya Ana Vásquez Ana Caballeros Carmen de Gutowski Lizeth Hernandex Brenda Barrios Brenda Moreira <u>Secondary</u> Rebeca de DeLeón Jose Ramirez Roberto Giron Alma Pérez Rodolfo Mendéz Francisco Guzman Guillermo Diéguez

COUNTRY	ORGANIZATION	PERSON(S) RESPONSIBLE FOR COMPLETING THE QUESTIONNAIRE	OTHERS CONSULTED
ISRAEL	National Institute for Testing and Evaluation	Dr Naomi Gafni	Dr Moshe Ilan Dr Yizchak Shapira Ms Ruth Raz Mr Gad Yair Ms Chayuta Regev Mrs Zehava Sass Dr Susan Fund
MALAYSIA	Malaysian Examinations Council	Mr Azumuddin Bin Mafizulla Mr Termuzi Bin Hj Abdul Aziz	Dr Dato 'Syed Abu Bakar Barakbah Mr Ahmed Bin Mohd Said Mr Mohd Rosli Bin Mohd Ghazali Mrs Teh Mun Hing Mr A Jayaratnam
NAMIBIA	University of Namibia, Faculty of Education	Professor Donton J Mkandawire	Mr P Koopman Mr J Erasmus Mr P Van der Merwe
POLAND	Ministerstwo Edukacji Narodowej	Professor Boleslaw Niemierko	Dr Jerzy Wisniewski Mrs Elzbieta Lamparska Mrs Iwona Michowska Mrs Bozena Krezel
SCOTLAND	The Scottish Examination Board	Dr David Elliot Mrs Christine de Luca Mrs Lucy Jones Mrs Lillian Monro	
SLOVENIA	National Examinations Centre	Dr Sergij Gaberscek	Mr Marjan Setinc

APPENDIX B

QUESTIONNAIRE

A COMPARATIVE STUDY OF CURRENT THEORIES AND PRACTICES IN ASSESSING STUDENTS' ACHIEVEMENTS AT PRIMARY AND SECONDARY LEVEL

INTRODUCTION

The purpose of the study is to examine and compare, in 10 selected countries, the interaction between current conceptions and theories of assessment and assessment practices.

The term 'assessment' is used throughout to include all forms of testing or examination of students' intellectual aptitudes, abilities and knowledge.

In particular the study seeks to identify in a comparative context

- the nature of changing practices and theories during the past 20-25 years in assessing the learning achievements of students in primary and secondary education;
- the socio-cultural and technical factors associated with the gaps between assessment theory and assessment practice;
- ways to reinforce links between theory and practice by increasing the relevance of the process of assessment to both local and national socio-cultural environments and the needs of individuals.

In responding to this questionnaire you are asked to:

- concentrate solely on assessment theory and practice in relation to students during the years of compulsory schooling and to exclude any developments which only effect students in further, higher, or vocational education;
- consider the full range of assessment practices, formal and informal, as they exist in primary and secondary education - for example: large scale public examinations, testing for selection or other purposes, procedures for marking and grading, everyday questioning and assessment strategies which are adopted in classrooms as a regular part of teaching and learning;
- identify by name and position those with whom you have consulted before presenting this data.

The questionnaire is in three sections:

- the first section asks for information about the nature of the school curriculum and assessment under four headings - the curriculum, statutory assessment which is legally required in schools, non-statutory formal assessment, classroom assessment;
- the second section invites you to evaluate current developments in assessment policies and practices within your educational system;
- the third section invites you to address any general issues not elsewhere covered in the questionnaire.

To facilitate later analysis you are asked to use our section and sub-section numbers to structure your response and index its presentation.

In this questionnaire compulsory schooling is assumed to consist of:

- primary education starting at or about the age of five and continuing until the age of 11; and
- secondary education for students aged 11-16.

If this does not apply in your country please describe the structure of the educational system in your response.

Please respond to all the questions in the first two sections twice: once with regard to primary education, and once with regard to the compulsory years of secondary education.

The first set of responses should be headed:

PRIMARY EDUCATION

SECTION P1 THE PRIMARY CURRICULUM AND ITS ASSESSMENT

SECTION P2 CURRENT DEVELOPMENTS (PRIMARY)

and the second set of responses should be headed:

SECONDARY EDUCATION

SECTION S1 THE SECONDARY CURRICULUM AND ITS ASSESSMENT

SECTION S2 CURRENT DEVELOPMENTS (SECONDARY)

The third set of responses should be headed:

SECTION III GENERAL ISSUES

Please address each question as fully and as factually as possible.

Where particular issues do not apply to your situation please indicate that this is so.

SECTION I : THE CURRICULUM AND ITS ASSESSMENT

1 THE CURRICULUM

1.1 Is the curriculum which students follow most appropriately described as

a largely common curriculum, or

a common core of subjects plus some optional subjects, or

a curriculum of wholly optional subjects?

1.2 Which subjects are common and which optional?

1.3 Is the curriculum wholly or partly determined by government legislation? If not, how are curricular decisions made?

1.4 Where there are privately run schools do they follow the same curriculum as other schools?

1.5 What allowance, if any, is made in the case of students who have particular curricular needs:

(i) those with physical/mental disabilities

(ii) those with learning difficulties

(iii) gifted students

(iv) those who attend schools where the teaching is not in their native language?

1.6 What elements, if any, have

(i) been added to the curriculum, and

(ii) been removed from the curriculum during the last 20-25 years and why?

ASSESSMENT

2. STATUTORY ASSESSMENT LEGALLY REQUIRED IN SCHOOLS

2.1 Is there an over-arching national and/or regional structure, established by law, for assessing students' achievements? (If the answer to this is no, please move to 3).

2.2 How long has it been in place?

2.3 For what purpose(s) was it introduced?

2.4 Does it involve assessment of the achievements of all students? If not - what is the basis on which exceptions are made (eg ability of student, type of institution attended)?

2.5 Which areas of the curriculum (academic, personal and social) are involved in the assessment process? How were these areas selected and others rejected?

2.6 What is the basis for the assessment structure?

(For example: the assessment model within the national curriculum for England and Wales is based on four criteria - that assessment results should be criterion-referenced; that they should be formative; that the scales or grades should be capable of comparison across classes and schools and should, therefore, be calibrated or moderated; and that the assessments should relate to progression across the eleven years of compulsory schooling).

2.7 How frequently are assessments made and at what age(s)?

2.8 Is the mode of the assessment tasks mainly:

(i) written

(ii) oral

(iii) practical

(iv) self assessment by students

(v) a combination of these?

2.9 Is the form of the written/oral assessment tasks mainly

(i) short answer

(ii) extended answer (eg essay)

(iii) multiple-choice

(iv) other?

2.10 Are the assessment tasks set internally (ie by students' teachers) or by an external body?

2.11 If assessment tasks are set by an external body, by whom is it appointed?

2.12 What steps are taken to ensure that tasks are free of economic, social or cultural bias?

2.13 Do all students of the same age undertake the same tasks? If not, how do the tasks differ and on what basis are different tasks given to students?

2.14 How are students' responses marked

(i) against a 'model' answer

(ii) against pre-determined and published criteria

(iii) using pre-determined norms for the whole population

(iv) application of standardised tests with answer keys

(v) using pooled judgements

(vi) moderated by consensus procedures

(vii) a combination of these (please specify)?

2.15 What steps are taken to ensure marking reliability?

2.16 How are the results achieved by students recorded

(i) as pass/fail

(ii) by scores

(iii) by grades

(iv) other (please specify)?

2.17 To whom are results reported (eg to individual students only, to parents, publicly)?

3. NON-STATUTORY FORMAL ASSESSMENT

(Non-statutory formal assessment refers to public examinations which might be taken by very large numbers of students although there is no legal requirement for them to do so. For example, in England and Wales, the GCSE examination would currently be one example of such formal assessment).

3.1 Excluding the kind of system identified in 2 above, what further assessment systems, if any, exist?

(For example: within the primary sector examinations to select students for different forms of secondary schooling; or, within the secondary sector, a qualifying examination for entry into further or higher education).

3.2 Describe any assessment system(s) you have identified in 3.1 above in terms of

(i) purpose

(ii) areas of the curriculum involved

(iii) mode of assessment (see 2.8)

(iv) form of assessment (see 2.9)

(v) setting of assessment tasks (see 2.10 and 2.11)

(vi) marking of students' responses (see 2.14)

(vii) recording results (see 2.16)

(viii) reporting results (see 2.17)

(ix) any other major features.

3.3 Does the assessment system(s) you have described in 3.2 in any way reflect social, economic or cultural differences between groupings of students?

(For example: are there different school selection examinations aimed at different social groups; how far do examinations assume a common/diverse cultural background?)

3.4 What innovations, if any, have there been within this assessment system(s) during the past 20-25 years?

3.5 Why have innovations taken place or, alternatively, why have the system(s) remained static?

3.6 Are the assessment systems you have identified valid (i.e. are they appropriate for the students at whom they are directed, do they assess what they claim to assess)?

4. CLASSROOM ASSESSMENT

(Classroom assessment refers to those formative assessment strategies which teachers use as an everyday part of their work in the classroom. It refers to assessment which is primarily designed to support and further students' learning.)

4.1 Are there national/local/institutional requirements or policies on the ways in which teachers should assess students' learning as a part of everyday classroom activity?

4.2 Do the requirements differ from one region to another or between institutions - if so in what way(s)?

4.3 Where there are differences are they mainly for academic, economic, social or cultural reasons?

4.4 Identify key ways in which students' learning is assessed in school by classroom teachers, from the following:

(i) application of standardised tests with answer keys

(ii) using criteria internalised by teachers

(iii) using pre-determined and published criteria

(iv) matching work to a model answer provided externally

(v) using pre-determined norms for the whole population

(vi) using pooled judgements

(vii) moderated by consensus procedures.

(viii) other practices.

4.5 What initial and/or in-service training on assessment is available to teachers?

4.6 How is the effectiveness of classroom assessment procedures monitored?

SECTION II : CURRENT DEVELOPMENTS

Considering in turn the primary and secondary sectors of your system

5.1 Comment on:

(i) conservative assessment practice

(ii) mainstream assessment practice

(iii) radical assessment practice

insofar as you are able to identify them. In each case describe the main features, estimate how widespread such practice is, where it is used and by whom.

5.2 How far, if at all, is there a gap between the stated purposes of the various assessment systems and the purposes for which they are used in practice (please refer, for example, to diagnostic and remedial purposes, guidance of teaching and learning, curricular evaluation, selection, appraisal of teachers and/or schools?)

5.3 How far do the existing assessment structures meet the needs of the contemporary society in which students live and of the future society in which they will be adults?
If there is a mis-match - why?

5.4 What, in your view, are the major changes which are needed in current assessment practices?

5.5 What pressures are there, from within the teaching profession or elsewhere, for changes to assessment practice?

SECTION III : GENERAL ISSUES

6. In this section you are invited to write as fully as you wish on any issues you feel are not covered, or are insufficiently addressed, elsewhere in the questionnaire.

Issues may include, for example:

socio-cultural factors operating in your country;

technical assessment problems;

meeting the individual needs of students;

marking and grading;

recording and reporting.

These are offered only as examples. They are not in any way intended to act as constraints.

BEST COPY AVAILABLE

Printed by Eyre & Spottiswoode Ltd. London and Margate